AKM Models

Matteo Paradisi
(EIEF and NBER)

Applied Micro - Lecture 8

Outline

We study two applications of fixed-effects and panel data

1. Value-added models
   - predict the impact of teachers, managers, judges
2. AKM Models
   - study the sources of wage inequality

Studying the Determinants of Wage Inequality

- Labor market outcomes driven by decisions of workers and firms
- Implications of workers and firms' heterogeneity for outcomes in the labor market
- matched employer-employee data allow analysis of both sides of the market
- Abowd, Kramarz and Margolis (1999) (AKM) started this literature

The Two-way FE Model

- Model with employer and employee heterogeneity, and observable covariates

$$y_{it} = \mu + x_{it}\beta + w_{jt}\gamma + u_i\eta + q_j\rho + \alpha_i + \phi_j + \epsilon_{it}$$

- $i = 1, \ldots, N$ workers
- $j = 1, \ldots, J$ firms
- y_{it} is typically wage
- x_{it}, u_i observable worker's covariates
- w_{jt}, q_j observable firm's covariates
- α_i, ϕ_j unobserved heterogeneity

The Two-way FE Model

$$y_{it} = \mu + x_{it}\beta + w_{jt}\gamma + u_i\eta + q_j\rho + \alpha_i + \phi_j + \epsilon_{it}$$

- Typically, we assume that α_i and ϕ_j are correlated with observables
- Hence, we cannot use random effects models
- We use FE models instead
- It follows that ρ and η cannot be identified
- Hence we define

$$\theta_i = \alpha_i + u_i\eta$$
$$\psi_j = \phi_j + q_j\rho$$

- And the model becomes

$$y_{it} = \mu + x_{it}\beta + w_{jt}\gamma + \theta_i + \psi_j + \epsilon_{it}$$

The Two-way FE Model

$$\theta_i = \alpha_i + \mathbf{u}_i \eta$$
$$\psi_j = \phi_j + \mathbf{q}_j \rho$$

- Under the assumption that $\text{Cov}(\mathbf{u}_i, \alpha_i) = \text{Cov}(\mathbf{q}_j, \phi_j) = 0$, we can identify ρ and η
- This allows to investigate how observables affect the time-constant heterogeneity of workers and firms
- However, often this is not a comfortable assumption
- Hence, this will not be the main focus of the analysis

Goals of the Two-Way FE Model

- What are our goals in this analysis?
- Study sorting of workers in firms
  - AKM title: "High wage workers in high wage firms"
  - Cov (θ_i, ψ_j) informative about sorting
- Study the determinants of the variance in wages

$$\text{Var}(y_{it}) = \text{Var}(\hat{\theta}_i) + \text{Var}(\psi_j) + 2\text{Cov}(\hat{\theta}_i, \hat{\psi}_j)$$

Identification

$$y_{it} = \mu + x_{it}\beta + w_{jt}\gamma + \theta_i + \psi_j + \epsilon_{it}$$

- The identification comes from movers only
- If workers do not move, we cannot identify the model: we cannot identify worker FE separately from firm
- Key assumption: mobility does not correlate with unobservables
- Important: we allow for sorting as long as it is NOT explained by time-varying unobservables

Identification - Formally

$$y_{it} = \mu + x_{it}\beta + w_{jt}\gamma + \theta_i + \psi_j + \epsilon_{it}$$

- Let's rewrite the model

$$y = Z\delta + D\theta + F\psi + \epsilon$$

- D: matrix of workers' dummy variables
- F: matrix of firm dummy variables
- $Z = (X, W)$: matrix of observables
- Key assumption

$$E\left(d_i'\varepsilon\right) = 0 \; \forall i$$
$$E\left(f_j'\varepsilon\right) = 0 \; \forall j$$

Identification - Formally

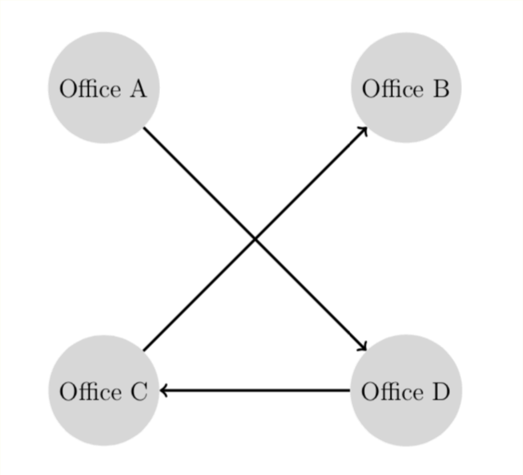
$$E\left(d_i'\varepsilon\right) = 0 \ \forall i$$
$$E\left(f_j'\varepsilon\right) = 0 \ \forall j$$

- As anticipated, workers can sort into firms
- BUT, sorting can only be explained by Zs (covariates) or by worker and firms FE
- Importantly, this specification does not allow for sorting based on competitive advantage a workers in particular firms
  - this would be an unobservable at the worker-firm level that correlates with d_i and f_j and that we cannot control for

Identification - Connected Sets

- The model exploits mobility for identification
- Hence, can identify workers and firms effects only in connected sets
- Set of firms that are indirectly connected by workers
- Because we estimate FE relative to excluded category, we cannot compare FE across connected sets
- There are some solutions though

Connected Set - Intuition



Source: Fenizia (2019)

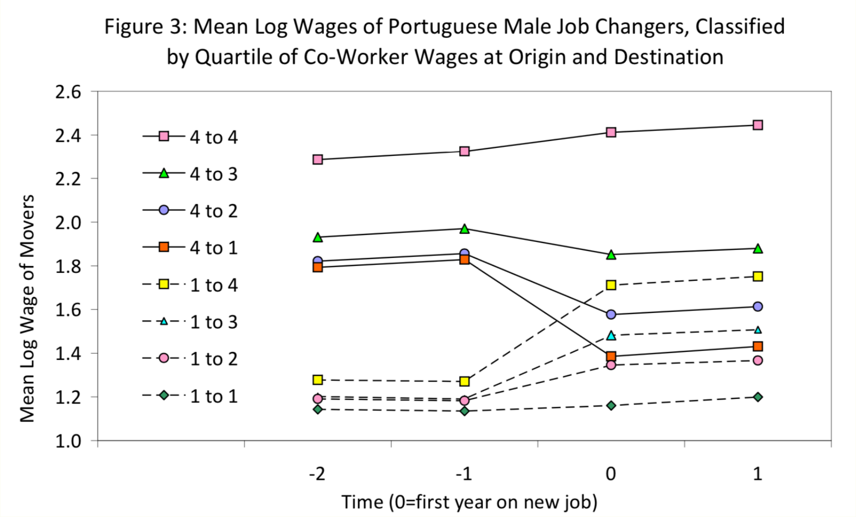
Identification - Implicit Restrictions of the Model

- FEs will be biased unless moves are uncorrelated with time-varying component of wages
- Hence, worker cannot move because of sudden drop in wages
- Card, Heining and Kline (2013) show a nice event study to investigate assumption
- Also, they use it to provide evidence of equal wage premium for all employees in same firm

Card, Heining and Kline (2013) Event Study

- They classify firms depending on quartile of the wage of co-workers
- Then look at wages around move date for any combination of moves
- Look for two things:
  1. if firms pay proportional premium to employees, we expect increase when moving to high quartile firm and decrease in the opposite case. Also, wage change in moving from A to B, must be symmetric to change of going from B to A
  2. If move does not depend on time-varying component of wage, then see flat wage trends before and after change

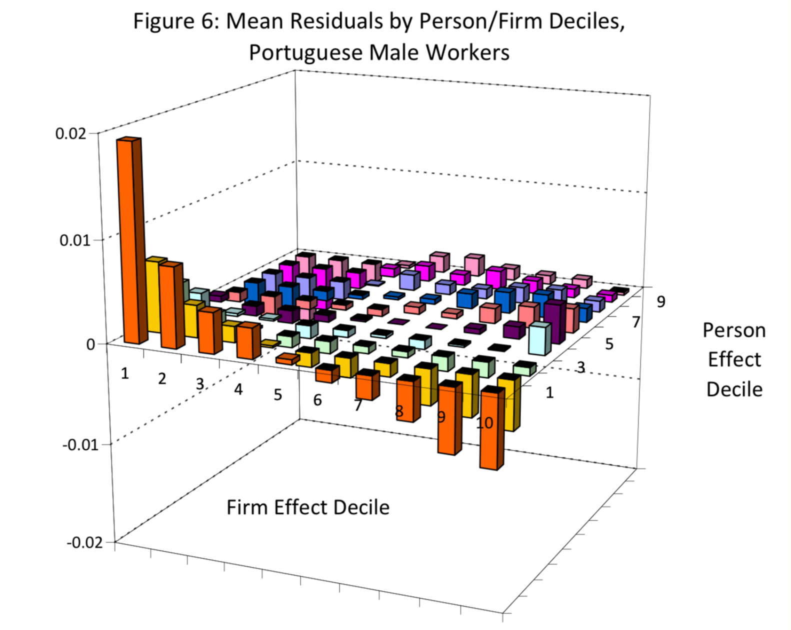
Card, Heining and Kline (2013) Event Study



Additivity Assumption

- Key assumption: all workers in same firm get the same wage premium
- This is a strong assumption, but there is a way to "test" it
- Look at residuals for different deciles of workers and firms FE
- If additivity holds, then we should see close to zero residuals

Checking Additivity



Checking Additivity

- Residuals are almost always close to zero
- Particularly, this is true for high-wage workers in high-wage firms
- Residuals differ from zero only for low-wage workers in low-wage firms
- One possible explanation is minimum wage

Estimating the Model

Estimating the Model

$$\mathbf{y} = \mathbf{Z}\delta + \mathbf{D}\theta + \mathbf{F}\psi + \epsilon$$

- You can think about using a dummy variable estimator
- In a model with only worker FE, we would include worker's dummy variables, which would be identical to within worker deviation from the mean
- BUT, in a model with two-way FEs we would need a double deviation from the mean
- However, this cannot be done in this model since it does not preserve the patterning in D and F
  - This is because there is no regular pattern in the mobility of workers across firms

Estimating the Model

$$y = Z\delta + D\theta + F\psi + \epsilon$$

- AKM solves the problem by including firm dummy variables, and taking deviations from workers' means
- This gives the same result as dummy variable estimator

$$M_D y = M_D Z\delta + M_D F\psi + M_D \epsilon$$

- $y_{it} - \bar{y}_i$ is regressed on covariates $z_{it} - \bar{z}_i$ and on J mean-deviated firm dummy vars $F^j_{it} - \bar{F}^j_i$
- This estimator is called FEiLSDVj (FE on i, dummy variable estimator on j)

Estimating the Model

- We obtain estimates of θ by inverting the equation

$$\mathsf{D}\hat{\theta} = \mathsf{P_D y} - \mathsf{P_D Z}\hat{\gamma} - \mathsf{P_D F}\hat{\psi}$$

- So for a single worker

$$\hat{\theta}_\mathsf{i} = \bar{\mathsf{y}}_\mathsf{i} - \bar{\mathsf{z}}_\mathsf{i}\hat{\gamma} - \mathsf{F}\bar{\hat{\psi}}_\mathsf{i}$$

- $\bar{\hat{\psi}}_\mathsf{i}$ is a weighted average of $\hat{\psi}_{\mathsf{j(it)}}$ over t, j (it) indicates the firm where worker i is a time j

- It follows that

$$\hat{\theta}_\mathsf{i} - \theta_\mathsf{i} = -\bar{\mathsf{z}}_\mathsf{i}\left(\hat{\gamma} - \gamma\right) - \left(\bar{\hat{\psi}}_\mathsf{i} - \bar{\psi}_\mathsf{i}\right) + \bar{\varepsilon}_\mathsf{i}$$

Estimating the Model

$$\hat{\theta}_i - \theta_i = -\bar{\mathbf{z}}_i\left(\hat{\gamma} - \gamma\right) - \left(\bar{\hat{\psi}}_i - \bar{\psi}_i\right) + \bar{\varepsilon}_i$$

- Hence, if conditional on z the ψ_j is overestimated, then on average the corresponding θ_i is underestimated
- The estimated correlation bw θ_i and ψ_j is underestimated, can we find a correction?

Estimates and a Puzzle

- AKM original paper reports a positive correlation bw θ and ψ: high-wage workers in high-wage firms
- We call this "assortative matching"
- However, subsequent papers find negative correlations
  - AKM (2004) in France
  - Gruetter and Lalive (2004) in Austria
  - Barth and Dale-Olsen (2003) in Norway
- Two possible reasons for this result
  - econometric motivation: error in estimates, bias correlation downward (see previous slides)
  - economic explanations
- Let's focus on econometrics

Bias in Variances and Covariance

- Andrews et al. (2008) derive closed-forms for the biases
- They show that

$$\mathsf{E}\left(\hat{\sigma}_\theta^2\right) = \sigma_\theta^2 + \text{Positive Bias}$$

$$\mathsf{E}\left(\hat{\sigma}_\psi^2\right) = \sigma_\psi^2 + \text{Positive Bias}$$

$$\mathsf{E}\left(\hat{\sigma}_{\theta\psi}\right) = \sigma_{\theta\psi}^2 + \text{Negative Bias}$$

- They propose a correction

Limited Mobility Bias

- Andrews et al. (2008) find a simple closed-form to illustrate "limited mobility bias"
- This bias had been discussed by AKM (2004) before
- They derive a formula for the simple case with balanced moves

$$\text{Bias} = -\frac{\sigma_{\varepsilon}^{2}}{N^{*}}\left(\frac{k}{M} - J\right)$$

- Keeping J fixed, number of moves M reduces the bias

Andrews et al. (2008) Correction

- Andrews et al. (2008) derive closed-forms for the biases
- Hence, they can use this closed forms for the biases to correct estimates of σs
- Strong assumption: errors are homoskedastic
- Can we relax the assumption?
- Anyway, they still find negative correlation

Kline Saggio and Solvsten (2019) Correction

- Kline et al. (2019) notice that to correct the bias in variances, one can use leave-one-out estimators
- Their procedure allows for heteroskedasticity in error terms
- They implement estimator on Italian data
- Show that correlation is positive

Kline Saggio and Solvsten (2019) Correction

Table 2: Variance Decomposition

	Pooled		**Younger Workers**		**Older Workers**	
	Leave one out sample	Leave two out sample	Leave one out sample	Leave two out sample	Leave one out sample	Leave two out sample
Variance of Log Wages	0.1843	0.1898	0.1200	0.1232	0.2591	0.2760
Variance of Firm Effects						
Plug in (PI)	0.0358	0.0316	0.0368	0.0314	0.0415	0.0304
Homoscedasticity Only (HO)	0.0295	0.0271	0.0270	0.0251	0.0350	0.0243
Leave Out (KSS)	0.0240	0.0238	0.0218	0.0221	0.0204	0.0180
Variance of Person Effects						
Plug in (PI)	0.1321	0.1341	0.0843	0.0827	0.2180	0.2406
Homoscedasticity Only (HO)	0.1173	0.1214	0.0647	0.0663	0.2046	0.2298
Leave Out (KSS)	0.1119	0.1179	0.0596	0.0634	0.1910	0.2221
Covariance Firm, Person Effects						
Plug in (PI)	0.0039	0.0077	-0.0058	-0.0008	-0.0032	-0.0006
Homoscedasticity Only (HO)	0.0097	0.0117	0.0030	0.0049	0.0040	0.0041
Leave Out (KSS)	0.0147	0.0149	0.0075	0.0075	0.0171	0.0115

New Developments

- Previous corrections are hard to implement with large datasets
- Bonhomme Lamandon and Manresa (2019) proposed a different strategy
- Their approach allows to account for rich patterns of complementarities and sorting
- They also build a dynamic version of the model where moves can be motivated by past earnings
- At the same time it solves the problem in variance estimation

Bonhomme Lamandon and Manresa (2019)

- Main intuition: reduce the heterogeneity in the problem
- They implement a two-step procedure
  1. Cluster firms based on their earnings distribution (using k-means clustering)
  2. Set up a likelihood function based on the move probabilities across the firms' clusters
- Because of clustering there is no problem of limited sample/mobility

Evidence of Strong Sorting

Variance decomposition ($\times 100$)

$\frac{Var(\alpha)}{Var(y)}$	$\frac{Var(\psi)}{Var(y)}$	$\frac{2Cov(\alpha,\psi)}{Var(y)}$	$\frac{Var(\varepsilon)}{Var(y)}$	$Corr(\alpha,\psi)$
60.03	2.56	12.17	25.24	49.13
(0.85)	(0.16)	(0.39)	(0.59)	(0.86)